Statistical methods in weather forecasting

Sándor Baran

University of Debrecen, Hungary

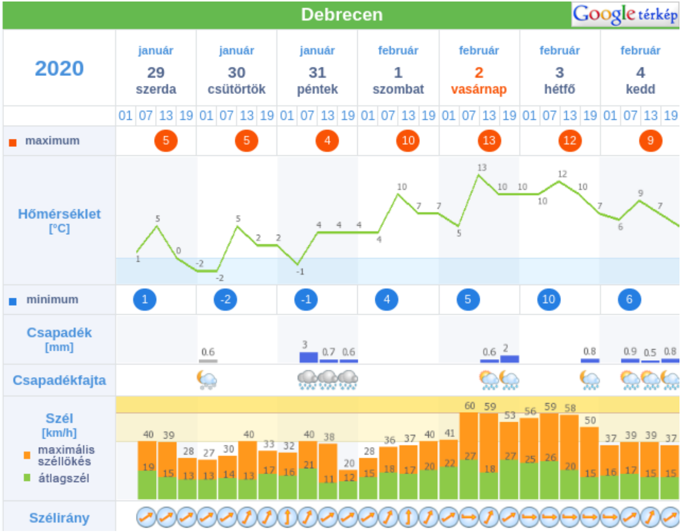
ICAI 2020, Eger

Research was supported by the Hungarian National Research, Development and Innovation Office under Grant No. NN125679 and by the EFOP-3.6.2-16-2017-00015 and EFOP-3.6.1-16-2016-00022 projects. The projects were co-financed by the Hungarian Government and the European Social Fund.

Outline

1. Probabilistic weather forecasting
2. Bayesian model averaging
3. Model training and model verification
4. Ensemble model output statistics
5. Case study: calibration of ALADIN-HUNEPS wind speed forecasts
6. Case study: calibration of water level forecasts
7. Further directions, ongoing projects

Point forecasts



www.met.hu

Point forecast: forecast obtained from a single run of a numerical weather prediction model[1].

A single forecast for each location, time and lead time.

Numerical solution of a complex system of non-linear partial differential equations. Initial conditions from observations.

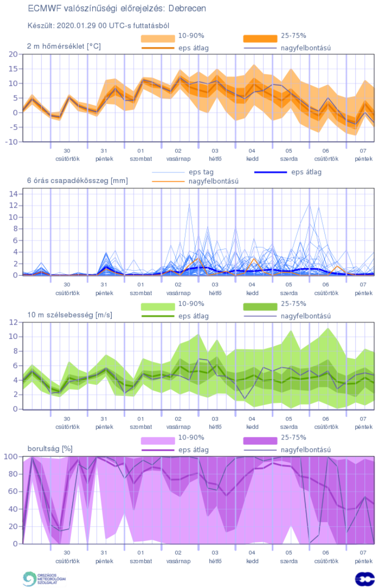
Advantage: easy to interpret.

Disadvantage: no information about uncertainty.

---

[1] Køltzov, M., Casati, B., Bazile, E., Haiden, T. and Valkonen, T. (2019) An NWP model intercomparison of surface weather parameters in the European Arctic during the Year of Polar Prediction Special Observing Period Northern Hemisphere 1. *Wea. Forecasting* **34**, 959–983.

Probabilistic forecasting



Ensemble of forecasts: forecasts obtained from multiple runs of a numerical weather prediction model with different initial conditions and/or model parametrization[1].

Several forecasts (8 – 255) for each location, time and lead time.

Quantities obtained from ensemble forecasts

- Point forecasts, e.g. ensemble mean, ensemble median.
- Dispersion, e.g. standard deviation.
- Probabilities of various events, e.g. surface (2 m) temperature is below 273.19 K (0 °C).
- Prediction intervals corresponding to a given level of confidence (e.g. 80%, 50%).

[1]Leith, C. E. (1974) Theoretical skill of Monte-Carlo forecasts. *Mon. Weather Rev.* **102**, 409–418.

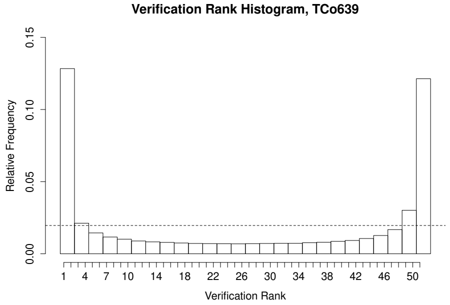
Ensemble forecasts

Recently all major meteorological services operate ensemble prediction systems:

- European Centre for Medium-Range Weather Forecasts: ECMWF ensemble, 51 members (since 1996);
- UK Met Office: MOGREPS ensemble, 18 members;
- US National Weather Service and Canadian Meteorological Center: NAEFS ensemble, 40 members;
- National Meteorological Center of CMA: T639 ensemble, 15 members;
- Deutscher Wetterdienst: COSMO-DE ensemble, 30 members;
- Hungarian Meteorological Service: ALADIN-HUNEPS ensemble, 11 members.

Problem: Ensemble forecasts often show an underdispersive character[1], or biased. Some form of post-processing is needed.

Example. Verification rank histograms of one-day ahead ECMWF global TCo639 temperature forecasts for JJA 2016.



[1] Buizza, R. (2018) Ensemble forecasting and the need for calibration. In Vannitsem, S., Wilks, D. S., Messner, J. W. (eds.), *Statistical Postprocessing of Ensemble Forecasts*, Elsevier, pp. 15–48.

Post-processing approaches

Aim: find the distribution of the future weather quantity usually with the help of ensembles and verifying observations[1].

Parametric models providing a full predictive distribution:

- Bayesian model averaging (BMA); `R` package `ensembleBMA`.

  Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* **133**, 1155–1174.

- Non-homogeneous regression or ensemble model output statistics (EMOS); `R` package `ensembleMOS`.

  Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118.

Data-based nonparametric approaches:

- Quantile regression: represents the predictive distribution by a finite set of its quantiles.

  Bremnes, J. B. (2004) Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Mon. Weather Rev.* **132**, 338–347.

- Member-by-member post-processing: transforms each of the ensemble members individually, results in a corrected ensemble.

  Van Schaeybroeck, B. and Vannitsem, S. (2015) Ensemble post-processing using member-by-member approaches: Theoretical aspects. *Q. J. R. Meteorol. Soc.* **141**, 807–818.

---

[1] Gneiting, T. and Raftery, A. E. (2005) Weather forecasting with ensemble methods. *Science* **310**, 248–249.

Post-processing approaches

Machine learning methods for deriving predictive distributions:

- Quantile regression forests: a generalization of random forests for quantile regression.

  Taillardat, M., Mestre, O., Zamo, M. and Naveau, P. (2016) Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Mon. Weather Rev.* **144**, 2375–2393.

- Neural networks
  - Regression estimation of parameters of predictive distributions.

    Rasp, S. and Lerch, S. (2018) Neural networks for postprocessing ensemble weather forecasts. *Mon. Weather Rev.* **146**, 3885–3900.
  - Quantile regression.

    Bremnes, J. B. (2020) Ensemble postprocessing using quantile function regression based on neural networks and Bernstein polynomials. *Mon. Weather Rev.* 148, 403–414.

- MLP, gradient boosting machines, random forests: estimation of discrete predictive distributions (e.g. total cloud cover), classification.

  Baran, Á, Lerch, S., El Ayari, M. and Baran, S. (2020) Machine learning for total cloud cover prediction. *arXiv:* 2001.05948.

Predictive distributions should be

- Calibrated: e.g. around 90% of the verifying observations should be contained between the lower and upper 5% quantiles of the predictive distributions.
- Sharp: e.g. the 90% central prediction intervals are narrower on average than the classical prediction intervals based on raw ensembles.

General Bayesian model averaging (BMA) model

X: weather quantity (vector) of interest (temperature, wind speed, wind vector, etc.).

$f_1, f_2, \ldots, f_K$: ensemble of forecasts for X with a given lead time.

$g_k(x \,|\, f_k; \theta_k)$: conditional PDF of X given f_k is the best forecast.

θ_k: parameter (vector) to be estimated with the help of training data.

BMA predictive PDF for X at a given location, time point and lead time[1]:

$$p\big(x \,\big|\, f_1, \ldots, f_K; \theta_1, \ldots, \theta_K\big) = \sum_{k=1}^{K} \omega_k g_k(x \,|\, f_k; \theta_k).$$

Mixture weights: $\omega_k \geq 0,\ k = 1, 2, \ldots, K,\ \omega_1 + \omega_2 + \ldots + \omega_K = 1.$

Exchangeable case: M ensemble members divided into K exchangeable groups[2].

Ensemble: $f_{k,\ell},\ k = 1, 2, \ldots, K,\ \ell = 1, 2, \ldots M_k.\ M_1 + M_2 + \ldots + M_K = M.$

$$p\big(x \,\big|\, f_{k,\ell}, k = 1, \ldots, K,\ \ell = 1, \ldots M_k; \theta_1, \ldots, \theta_K\big) = \sum_{k=1}^{K} \sum_{\ell=1}^{M_k} \omega_k g_k(x \,|\, f_{k,\ell}; \theta_k).$$

[1] Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118.

[2] Fraley, C., Raftery, A. E. and Gneiting, T. (2010) Calibrating multimodel forecast ensembles with exchangeable and missing members using Bayesian model averaging. *Mon. Weather Rev.* **138**, 190–202.

Normal BMA model

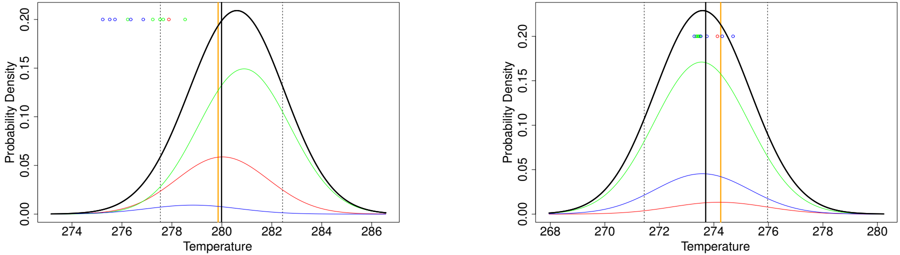
Predictive distribution for temperature or sea level pressure:

$$\sum_{k=1}^{K} \omega_k \mathcal{N}(b_{k,0} + b_{k,1} f_k, \sigma^2).$$

Parameters to be estimated: $b_{k,0}, \ b_{k,1}, \ \omega_k, \ k = 1, 2, \ldots, K,$ and σ^2.

Estimation of $b_{0k}, \ b_{1k}, \ k = 1, 2, \ldots, K$: linear regression; $\omega_1, \ldots, \omega_K$ and σ^2: ML with EM algorithm.

Example. ALADIN-HUNEPS temperature forecasts for Debrecen Airport[1]



24 November 2010 21 December 2010

[1] Baran, S., Horányi, A. and Nemoda, D. (2014) Probabilistic temperature forecasting with statistical calibration in Hungary. *Meteorol. Atmos. Phys.* **124**, 129–142.

Further univariate BMA models

Wind speed

- Gamma mixture.

  Sloughter, J. M., Gneiting, T. and Raftery, A. E. (2010) Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *J. Am. Stat. Assoc.* **105**, 25–37.

- Truncated normal mixture with cut-off at zero from below.

  Baran, S. (2014) Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components. *Comput. Stat. Data. Anal.* **75**, 227–238.

Precipitation accumulation

- Discrete-continuous model. Point mass at zero, gamma mixture for modelling positive precipitation accumulation.

  Sloughter, J. M., Raftery, A. E., Gneiting, T. and Fraley, C. (2007) Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Mon. Weather Rev.* **135**, 3209–3220.

Wind direction

- Mixture of Von-Mises distributions.

  Bao, L., Gneiting, T., Raftery, A. E., Grimit, E. P. and Guttorp, P. (2010) Bias correction and Bayesian model averaging for ensemble forecasts of surface wind direction. *Mon. Weather Rev.* **138**, 1811–1821.

Hydrological BMA models

Box-Cox transformed streamflow, water level

- Normal mixture.

  Duan, Q., Ajami, N. K., Gao, X. and Sorooshian, S. (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.* **30**, 1371–1386.

- Doubly truncated normal mixture.

  Baran, S., Hemri, S. and El Ayari, M. (2019) Statistical post-processing of water level forecasts using Bayesian model averaging with doubly-truncated normal components. *Water Resour. Res.* **55**, 3997–4013.

Bivariate models

Wind vector

- Bivariate normal mixture.

  Sloughter, J. M., Gneiting, T. and Raftery, A. E. (2013) Probabilistic wind vector forecasting using ensembles and Bayesian model averaging. *Mon. Weather Rev.* **141**, 2107–2119.

Wind speed and temperature

- Bivariate normal mixture truncated from below at zero in the wind coordinate.

  Baran, S. and Möller, A. (2015) Joint probabilistic forecasting of wind speed and temperature using Bayesian model averaging. *Environmetrics* **26**, 120–132.

Training data

Training data: a set of ensemble forecasts and verifying observations.

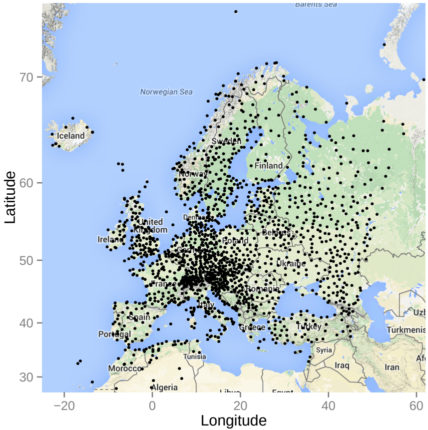
Rolling training period: a sliding window with data from the preceding n time points.

Regional estimation: parameters are estimated using all available forecast cases from the training period. A single universal set of parameters across the entire ensemble domain. Short training periods.

Local estimation: distinct parameter estimates for the different stations using only the training data of the given station.

Semi-local estimation: parameters for a given station are estimated using training data of similar stations[1].

- Distance-based similarity.
- Clustering-based similarity.



Domain of the GLAMEPS ensemble

[1] Lerch, S. and Baran, S. (2017) Similarity-based semi-local estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Stat.* **66**, 29–51.

Distance based similarity

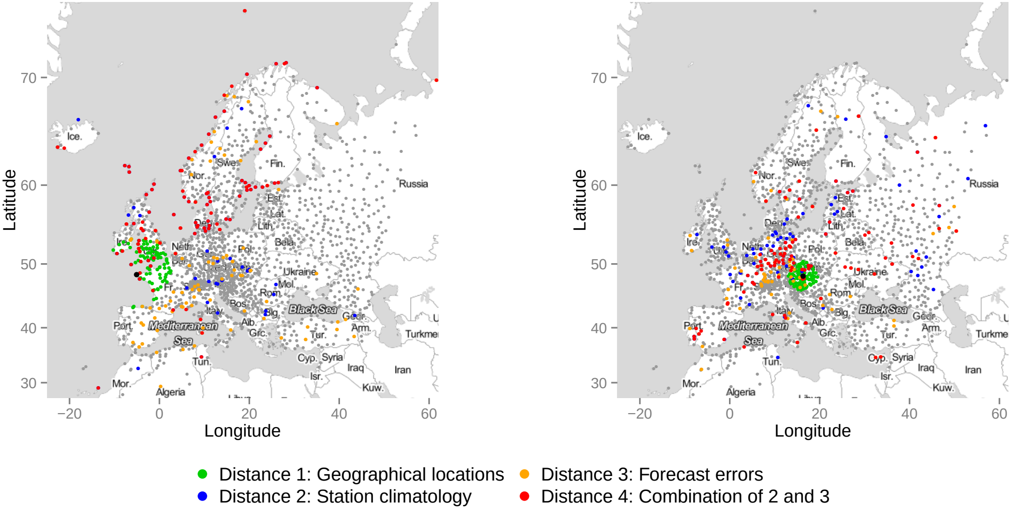


Illustration of the 100 most similar stations measured by the four distance functions for two reference stations at Ouessant, France (*left*) and Vienna, Austria (*right*).

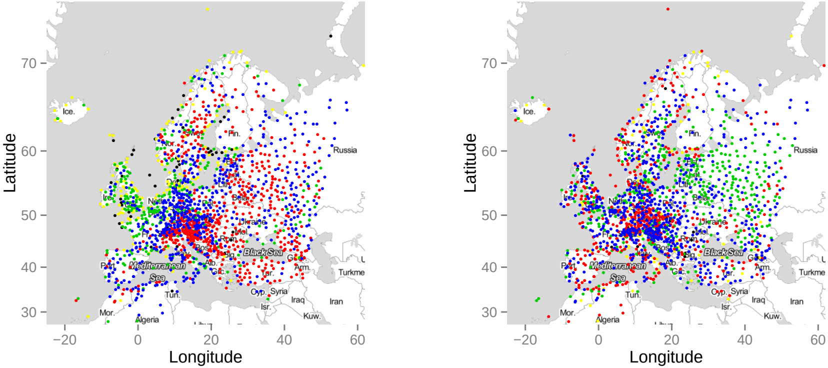
Clustering

k-means clustering, each station is characterized by an N-dimensional feature vector.

1. Station climatology. Equidistant $\frac{1}{N+1}, \frac{2}{N+1}, \ldots, \frac{N}{N+1}$ quantiles of the empirical CDF of observations over the training period.

2. Forecast errors. Equidistant quantiles of the empirical CDF of the forecast errors of the ensemble mean.

3. Combination of feature sets 1 and 2. 50 - 50 %.



Clustering based on feature sets 1. (*left*) and 2. (*right*). 5 clusters, 24 features.

Verification scores

$x_{s,t}$: observation of the weather or hydrological quantity at location s and time t.

$p_{s,t}(x)$, $P_{s,t}(x)$: estimated predictive PDF and CDF at location s and time t with a given lead time.

$\widehat{x}_{s,t}$: point forecast based on $p_{s,t}(x)$ or on the ensemble (mean, median).

Deterministic forecasts:

- MAE: $\frac{1}{n}\sum_{s,t}\left|x_{s,t}-\widehat{x}_{s,t}\right|$; RMSE: $\sqrt{\frac{1}{n}\sum_{s,t}\left(x_{s,t}-\widehat{x}_{s,t}\right)^2}$, where n is the total number of forecast cases.

Probabilistic forecasts:

- Mean continuous ranked probability score (CRPS):

$$\overline{\mathrm{CRPS}} := \frac{1}{n}\sum_{s,t}\mathrm{CRPS}\left(P_{s,t}, x_{s,t}\right), \quad \text{with} \quad \mathrm{CRPS}\left(P, x\right) := \int_{-\infty}^{\infty}\left(P(y)-\mathbb{1}_{\{y\geq x\}}\right)^2 \mathrm{d}y.$$

  Improvement in CRPS with respect to a reference predictive distribution P_{ref}: continuous ranked probability skill score (CRPSS)

$$\mathrm{CRPSS} := 1-\frac{\overline{\mathrm{CRPS}}}{\overline{\mathrm{CRPS}}_{ref}}.$$

- Mean logarithmic score: $\overline{\mathrm{LogS}} := \frac{1}{n}\sum_{s,t}\Big(-\log\left[p_{s,t}(x_{s,t})\right]\Big)$.
- Coverage: percentage of observations in, e.g. 90 % central prediction interval.
- Sharpness: average width of central prediction interval.

Ensemble model output statistics (EMOS)

Predictive PDF is a single parametric density where the parameters are functions of the ensemble. Parameters of these functions are estimated by optimizing the average value of some verification score over the training data.

Temperature and pressure

Predictive distribution for temperature or sea level pressure:

$$\mathcal{N}\big(a_0 + a_1 f_1 + \cdots + a_K f_K, b_0 + b_1 S^2\big) \qquad \text{with} \qquad S^2 := \frac{1}{K-1}\sum_{k=1}^{K}\big(f_k - \overline{f}\big)^2.$$

Parameters to be estimated: $a_0, \ a_1, \dots, a_K \in \mathbb{R}$ and $b_0, b_1 \geq 0$.

Minimum points of the mean value of an appropriate verification score over the training data. Usually mean CRPS or mean logarithmic score. CRPS can be given in a closed form.

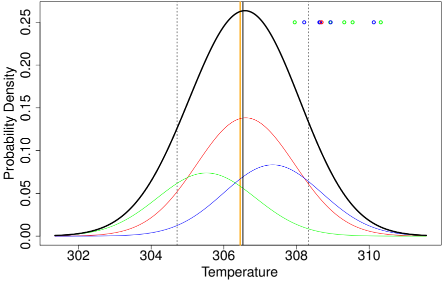
Exchangeable ensemble members:

$$\mathcal{N}\Bigg(a_0 + a_1 \sum_{\ell_1=1}^{M_1} f_{1,\ell_1} + \cdots + a_m \sum_{\ell_K=1}^{M_K} f_{K,\ell_K}, b_0 + b_1 S^2\Bigg).$$

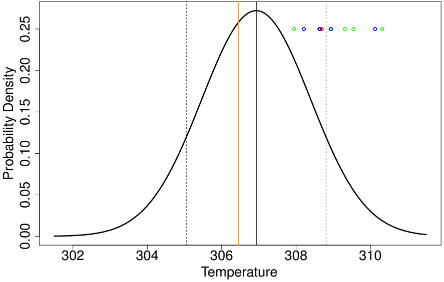
Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118.

Example

ALADIN-HUNEPS temperature forecasts and predictive distributions for Debrecen Airport for 2 July 2012.



BMA



EMOS

Baran, S., Horányi, A. and Nemoda, D. (2014) Comparison of the BMA and EMOS statistical methods in calibrating temperature and wind speed forecast ensembles. *Időjárás* **118**, 217–241.

Further univariate EMOS models

Wind speed

- Truncated normal (TN) distribution with cut-off at zero from below.

  Thorarinsdottir, T. L. and Gneiting, T. (2010) Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *J. R. Stat. Soc. Ser. A Stat. Soc.* **173**, 371–388.
- Generalized extreme value (GEV) distribution and TN-GEV regime switching.

  Lerch, S. and Thorarinsdottir, T. L. (2013) Comparison of non-homogeneous regression models for probabilistic wind speed forecasting. *Tellus A* **65**, 21206.
- Log-normal (LN) distribution and TN-LN regime switching.

  Baran, S. and Lerch, S. (2015) Log-normal distribution based EMOS models for probabilistic wind speed forecasting. *Q. J. R. Meteorol. Soc.* **141**, 2289–2299.
- TN-LN mixture distribution.

  Baran, S. and Lerch, S. (2016) Mixture EMOS model for calibrating ensemble forecasts of wind speed. *Environmetrics* **27**, 116–130.

Precipitation accumulation

- Censored generalized extreme value distribution.

  Scheuerer, M. (2014) Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Q. J. R. Meteorol. Soc.* **140**, 1086–1096.
- Censored, shifted gamma distribution.

  Baran, S. and Nemoda, D. (2016) Censored and shifted gamma distribution based EMOS model for probabilistic quantitative precipitation forecasting. *Environmetrics* **27**, 280–292.

Hydrological EMOS models

Box-Cox transformed water runoff, water level

- Left censored (and right truncated) normal distribution.

  Hemri, S., Lisniak, D. and Klein, B. (2014) Ermittlung probabilistischer Abflussvorhersagen unter Berücksichtigung zensierter Daten. *HyWa* **58**, 84–94.

- Right truncated normal distribution.

  Hemri, S., Lisniak, D. and Klein, B. (2015) Multivariate postprocessing techniques for probabilistic hydrological forecasting. *Water Resour. Res.* **51**, 7436–7451.

- Doubly truncated normal distribution.

  Hemri, S. and Klein, B. (2017) Analog based post-processing of navigation-related hydrological ensemble forecasts. *Water Resour. Res.* **53**, 9059–9077.

Bivariate EMOS models

Wind vector

- Bivariate normal distribution.

  Schuhen, N., Thorarinsdottir, T. L. and Gneiting, T. (2012) Ensemble model output statistics for wind vectors. *Mon. Weather Rev.* **140**, 3204–3219.

Wind speed and temperature

- Bivariate normal distribution truncated from below at zero in the wind coordinate.

  Baran, S. and Möller, A. (2017) Bivariate ensemble model output statistics approach for joint forecasting of wind speed and temperature. *Meteorol. Atmos. Phys.* **129**, 99–112.

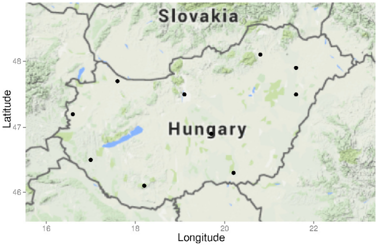
ALADIN-HUNEPS ensemble

Wind speed data of 10 major cities (Miskolc, Szombathely, Győr, Budapest, Debrecen, Nyíregyháza, Nagykanizsa, Pécs, Kecskemét, Szeged) of Hungary. Source: HMS.

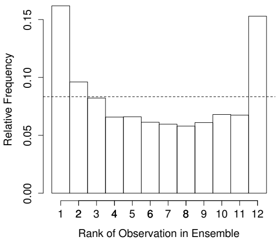
Data: 42h forecasts of 10m instantaneous wind speed (m/s) produced by the ALADIN-HUNEPS system of the HMS and the corresponding validating observations.

Ensemble: 10 exchangeable and one control forecast initialized at 18 UTC. Forecasts for 12 UTC two days later.

Period: 1 April 2012 – 31 March 2013. Missing: 6 days (excluded).



Verification Rank Histogram



Calibration of raw forecasts

Ensemble range contains the observed wind speed only in 61.21 % of the cases.

Underdispersive and uncalibrated. Nominal coverage: 83.33 %.

Horányi, A., Mile, M. and Szűcs, M. (2011) Latest developments around the ALADIN operational short-range ensemble prediction system in Hungary. *Tellus A* **63**, 642–651.

Calibration models for ALADIN-HUNEPS wind speed forecasts

Two groups: control member and 10 exchangeable members.

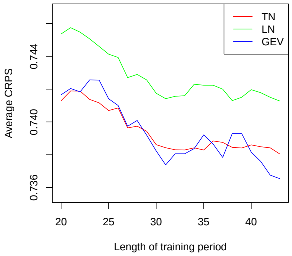
Post-processing models:

- Gamma BMA model
- Truncated normal BMA model
- Truncated normal (TN) EMOS model
- Log-normal (LN) EMOS model
- Generalized extreme value (GEV) EMOS model
- TN-LN regime-switching EMOS model.
  Ensemble median below a threshold: TN EMOS, otherwise: LN EMOS.
- TN-GEV regime-switching EMOS model
  Ensemble median below a threshold: TN EMOS, otherwise: GEV EMOS.
- TN-LN mixture EMOS model. Weighted mixture of TN and LN distributions.
  Optimum score estimation with respect to mean CRPS and mean logarithmic score.

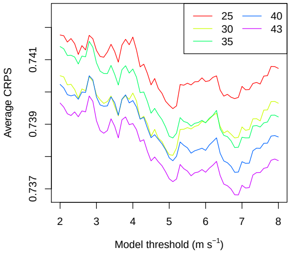
Training: Global parameter estimation, rolling training period.

Optimal training period length and thresholds

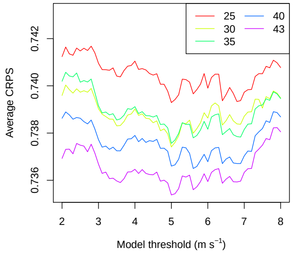
Studies based on CRPS, MAE and RMSE scores of BMA and TN EMOS models: optimal training period length is 43 days[1].



CRPS vs. training period length



CRPS vs. threshold, TN-LN



CRPS vs. threshold, TN-GEV

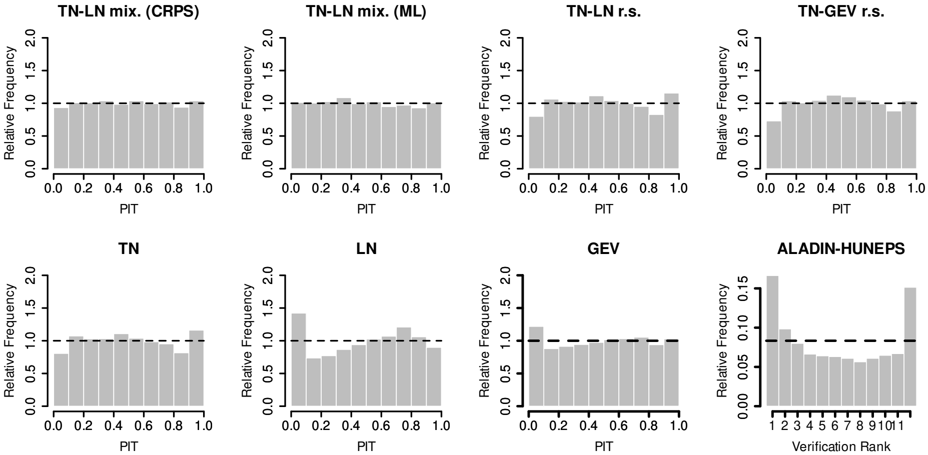
43 days training period length can be kept. BMA and EMOS predictive PDFs and validating observations for 313 calendar days (3130 forecast cases).

Optimal threshold values: TN-LN $\theta = 6.9\,\mathrm{m/s}$; TN-GEV $\theta = 5.0\,\mathrm{m/s}$.

LN proportion in TN-LN model: 4 %; GEV proportion in TN-GEV model: 15 %.

[1] Baran, S., Horányi, A. and Nemoda, D. (2014) Comparison of the BMA and EMOS statistical methods in calibrating temperature and wind speed forecast ensembles. *Időjárás* **118**, 217–241.

PIT histograms



Bootstrap estimates of rejection rates of the α^0_{1234} test of uniformity based on 10 000 random samples of size 2 500 each at the 0.05 level.

TN-LN mix. (CRPS)	TN-LN mix. (ML)	TN-LN r.s.	TN-GEV r.s.	TN	LN	GEV
0 %	1 %	100 %	100 %	100 %	100 %	79 %

Verification scores

Mean CRPS of probabilistic, MAE (median) and RMSE (mean) of point forecasts (m/s), coverage (%) and average width (m/s) of 83.33 % central predictions intervals.

Predictive model	CRPS	MAE	RMSE	Coverage	Av. w.
TN	0.738	1.037	1.357	83.59	3.53
LN	0.741	1.038	1.362	80.44	3.57
TN-LN mix. (CRPS)	0.736	1.037	1.358	83.02	3.62
TN-LN mix. (ML)	0.737	1.040	1.360	83.14	3.58
TN-LN r.s., $\theta = 6.9$	0.737	1.035	1.356	83.59	3.54
GEV	0.737	1.041	1.355	81.21	3.54
TN-GEV r.s., $\theta = 5.0$	0.735	1.039	1.355	85.59	3.72
Gamma BMA	0.760	**1.075**	**1.427**	81.87	3.72
TN BMA	0.698	1.045	1.377	85.46	3.76
Ensemble	0.803	1.069	1.373	68.22	2.88
Climatology	1.046	1.481	1.922	82.54	3.43

- Differences between the CRPS values of mixture and regime-switching models are not significant.
- Mean (maximal) probability of forecasting a negative wind speed:
  GEV model 0.33 % (9.46 %); TN-GEV model 2.74×10^{-3} % (0.15 %).

Hydrological ensemble forecasts

Hydrological ensemble forecasts are obtained by plugging in the ensemble weather forecasts into hydrological models. Water levels are

- bounded both from above and below;
- typically non-Gaussian, so Box-Cox transformation is applied[1] to normalize them.

Source: icis.com

Data: Water levels at Kaub gauge of the Rhine river and 79-member multimodel ensemble forecasts with lead times 1-120h initialized at 6 UTC for the period 1 January 2008 – 31 December 2015[2].

- 1 member from ECMWF high-resolution;
- 51 members from ECMWF ensemble;
- 16 members from COSMO LEPS;
- 11 members from NCEP GEFS ensemble.

[1] Duan, Q., Ajami, N. K., Gao, X. and Sorooshian, S. (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.* **30**, 1371–1386.

[2] Hemri, S. and Klein, B. (2017) Analog based post-processing of navigation-related hydrological ensemble forecasts. *Water Resour. Res.* **53**, 9059–9077.

Models for water level forecasts at gauge Kaub

Model formulation: 4 groups of exchangeable ensemble members according to the ensemble components.

Post-processing methods: doubly truncated normal BMA[1] and EMOS[2].

Free parameters to be estimated: 12 for BMA model (2 parameter estimation methods); 7 for EMOS model.

Training: 100-day rolling training period.

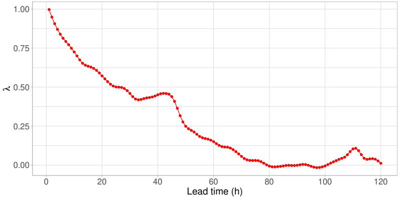
Verification: 10 April 2008 – 31 December 2015 (2822 calendar days).

Bounds: half of the minimum and double of the maximum recorded water level.
Lower bound: 16.5 cm; upper bound: 1650 cm.

Box-Cox transformation:

$$h_\lambda(x) := \begin{cases} (x^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \log(x), & \lambda = 0. \end{cases}$$

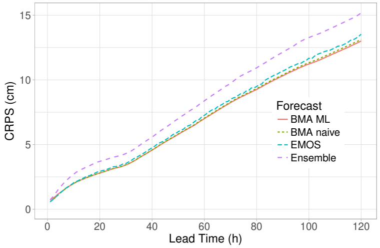
Different parameter λ for each lead time providing the best fit of the EMOS models[2].

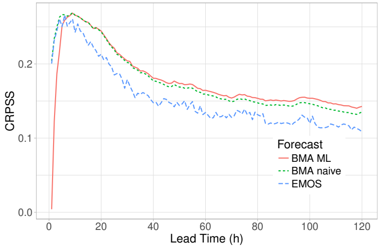


[1] Baran, S., Hemri, S. and El Ayari, M. (2019) Statistical post-processing of hydrological forecasts using Bayesian model averaging. *Water Resour. Res.* **55**, 3997–4013.

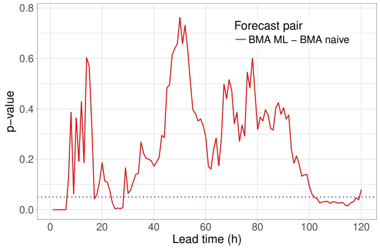
[2] Hemri, S. and Klein, B. (2017) Analog based post-processing of navigation-related hydrological ensemble forecasts. *Water Resour. Res.* **53**, 9059–9077.

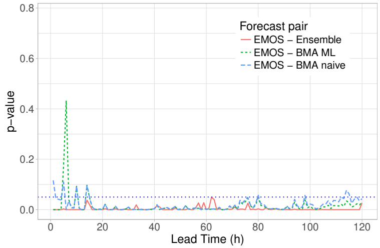
Continuous ranked probability score





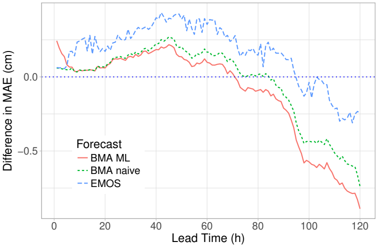
Mean CRPS values (*left*) and CRPSS with respect to the raw ensemble (*right*).

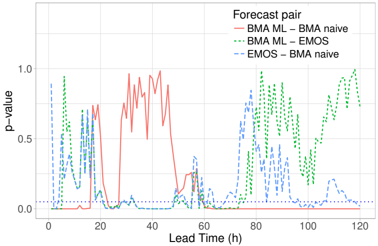




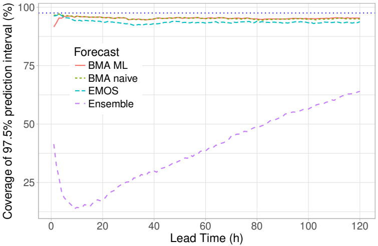
p-values of Diebold-Mariano tests for equality of mean CRPS of the two BMA approaches (*left*) and of all models compared to EMOS (*right*).

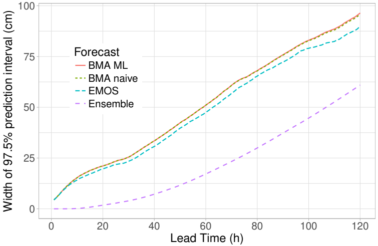
Mean absolute error, coverage, sharpness





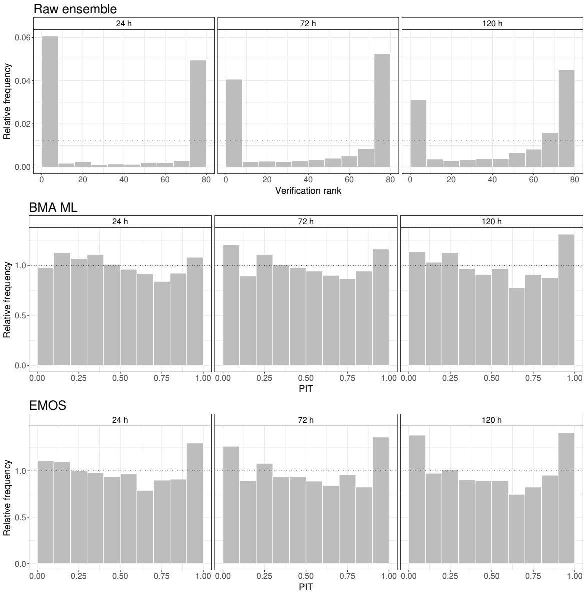
Difference in MAE values from the raw ensemble (*left*); *p*-values of Diebold-Mariano tests for equality of MAE (*right*).





Coverage (*left*) and average width (*right*) of nominal 97.5 % central prediction intervals.

Verification rank histogram, PIT histograms



Further directions of research, ongoing projects

- Application of post-processing to various ensemble prediction systems (ongoing work).

  Díaz, M., Nicolis, O., Marín, J. C. and Baran, S. (2020) Statistical post-processing of ensemble forecasts of temperature in Santiago de Chile. *Meteorol. Appl.* **27**, paper e1818.

- Calibration of ECMWF dual-resolution temperature and precipitation ensemble forecasts (ongoing work).

  Baran, S., Leutbecher, M., Szabó, M. and Ben Bouallègue, Z. (2019) Statistical post-processing of dual-resolution ensemble forecasts. *Q. J. R. Meteorol. Soc.* **145**, 1705–1720.

- Model development for heat indices (ongoing work).

  Baran, S., Baran, Á., Pappenberger, F. and Ben Bouallègue, Z. (2020) Statistical post-processing of heat index ensemble forecasts: is there a royal road? *arXiv:* 2001.08712.

Partners:

- European Centre for Medium-Range Weather Forecasts;
- Heidelberg Institute for Theoretical Studies;
- Hungarian Meteorological Service;
- University of Valparaíso.